Application 09/161,770
GAU: 2721

Patent Office
Canberra

RECEIVED
MAY 3 1999
Group 2700

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES, hereby certify that the annexed is a true copy of the Provisional specification in connection with Application No. PO 9510 for a patent by CANON KABUSHIKI KAISHA and CANON INFORMATION SYSTEMS RESEARCH AUSTRALIA PTY LTD filed on 29 September 1997.

I further certify that the annexed specification is not, as yet, open to public inspection.

RECEIVED
NOV 09 1998
Group 2700

WITNESS my hand this First
day of October 1998

KIM MARSHALL
MANAGER EXAMINATION SUPPORT AND SALES

S & F Ref: 395592

ORIGINAL

AUSTRALIA

Patents Act 1990

PROVISIONAL SPECIFICATION FOR THE INVENTION ENTITLED:

A Method of Data Compression

Name and Address of Applicant: Canon Kabushiki Kaisha, incorporated in Japan, of 30-2, Shimomaruko 3-chome, Ohta-ku, Tokyo, 146, JAPAN; Canon Information Systems Research Australia Pty Ltd, an Australian Company ACN No. 003 943 780, of 1 Thomas Holt Drive, North Ryde, New South Wales, 2113, AUSTRALIA

This invention is best described in the following statement:

5805

A Method for Digital Data Compression

Field of the Invention

The present invention relates to the field of data compression with particular application to digital image compression. More particularly, the present invention discloses a digital image compression method using spatial magnitude context entropy coding of discrete wavelet transform coefficients.

Background of the Invention

The field of digital data compression and in particular digital image compression has attracted great interest for some time.

In the field of digital image compression, many different techniques have been utilised. In particular, one popular technique is the JPEG standard which utilises the discrete cosine transform to transform standard size blocks of an image into corresponding cosine components. In this respect, the higher frequency cosine components are heavily quantised so as to assist in obtaining substantial compression factors. The heavy quantisation is an example of a "lossy" technique of image compression. The JPEG standard also provides for the subsequent lossless compression of the transformed coefficients.

Recently, the field of wavelet transforms has gained great attention as an alternative form of data compression. The wavelet transform has been found to be highly suitable in representing data having discontinuities such as sharp edges. Such discontinuities are often present in image data or the like.

Although the preferred embodiments of the present invention will be described with reference to the compression of image data, it will be readily evident that the preferred embodiment is not limited thereto. For examples of the many different applications of Wavelet analysis to signals, reference is made to a survey article entitled "Wavelet Analysis" by Bruce et. al. appearing in

IEEE Spectrum, October 1996 page 26 - 35. For a discussion of the different applications of wavelets in computer graphics, reference is made to "Wavelets for Computer Graphics", I. Stollinitz et. al. published 1996 by Morgan Kaufmann Publishers, Inc.

Summary of the Invention

It is an object of the present invention to provide for the efficient form of representation of data.

In accordance with the first aspect of the present invention there is disclosed a method of compressing data comprising applying a transform to the data to produce transformed data having a series of parts; entropy encoding the magnitude of the transformed data of at least one of said parts; and separately encoding the value of said transformed data. Preferably said entropy encoding utilizes the number of non-zero coefficients surrounding a spatial location of a corresponding transformed data value. The entropy encoding can comprise encoding the number of leading zeros in transformed data values.

The method further preferably comprises quantizing transformed portions of said data to integer values including a sign bit and a predetermined number of coefficient bits. Ideally, the preferred embodiment includes wavelet transforming the data with each of the sub-band components of the wavelet transform being separately entropy encoded. The present invention is ideally suited to the compression of image data and can be implemented on a standard computer device.

Brief Description of the Drawings

Notwithstanding any other forms which may fall within the scope of the present invention, preferred forms of the invention will now be described, by way of example only, with reference to the accompanying drawings in which:

Figs. 1-3 illustrate the process of wavelet transforming image data;

Fig. 4 illustrates the steps involved in the encoder of

the preferred embodiment;

Fig. 5 illustrates the steps in the decoder as constructed in accordance with the preferred embodiment;

Fig. 6 illustrates a data structure utilised by the preferred embodiment;

Fig. 7 illustrates the process of utilising a surrounding context for a current coefficient; and

Fig. 8 illustrates the process of histogram calculation for each subband.

Description of Preferred and Other Embodiments

The preferred embodiment proceeds initially by means of a wavelet transform of image data. A description of the wavelet transform process is given in many standard texts and in particular the aforementioned book by Stollnitz et. al. An overview of the wavelet process will now be described with reference to the accompanying drawings.

Referring initially to Fig. 1, an original image 1 is transformed utilising a Discrete Wavelet Transform (DWT) into four subimages 3-6. The subimages or subbands are normally denoted LL1, HL1, LH1 and HH1. The one suffix on the subband names indicates level 1. The LL1 subband is a low pass decimated version of the original image.

The wavelet transform utilised can vary and can include, for example, Haar basis functions, Daubechies basis functions etc. The LL1 subband is then in turn utilised and a second Discrete Wavelet Transform is applied as shown in Fig. 2 giving subbands LL2 (8), HL2 (9), LH2 (10), HH2 (11). This process is continued for example as illustrated in Fig. 3 wherein the LL4 subband is illustrated, the LL4 band decomposition process being referred to as an octave band filter bank with the LL4 subband being referred to as the DC subband. Obviously, further levels of decomposition can be provided depending on the size of the input image.

Each single level DWT can in turn be inverted to obtain the original image. Thus a J-level DWT can be inverted as a series of J-single level inverse DWT's.

To code an image hierarchically the DC subband is coded first. Then, the remaining subbands are coded in order of decreasing level. That is for a 4 level DWT, the subbands at level 4 are coded after the DC subband (LL4). That is the HL4, LH4 and HH4 subbands. The subbands at level 3 (HL3, LH3, and HH3) are then coded, followed by those at level 2 (HL2, LH2 and HH2) and then level 1 (HL1, LH1 and HH1).

With standard images, the encoded subbands normally contain the "detail" information in an image. Hence, they often consist of a sparse array of values and substantial compression can be achieved by quantisation of the subbands and efficient encoding of their sparse matrix form.

In the preferred embodiment, an effective compression of the subbands is provided through the utilisation of the correlation between the energy of adjacent coefficients in the DWT subband.

The encoding proceeds via two stages. The first stage encodes the leading zeros, or magnitude category of each DWT coefficient, based on the context of the number of surrounding coefficients that are not quantised to zero. By using a limited window this number takes on a relatively small range of values, and hence there are a limited number of contexts. For efficient entropy coding such a small number of contexts is desired. After the number of leading zero is transmitted or encoded the remaining bits for coefficients that are not quantised to zero are transmitted or encoded as is. Entropy coding could also be used at this stage. However, experiments suggest that this may not substantially increase the compression, while introducing more complexity.

An overview of the coding process is illustrated 20 in Fig. 4, while the decoding process is illustrated 30 in Fig. 5.

Turning initially to Fig. 4, a digital image is transformed 21 using a Discrete Wavelet Transform into

several subband components as previously described. Each subband is preferably coded in a hierarchical order. As illustrated in Fig. 6, each coefficient in a subband is quantised to an integer value 25 having a predetermined number of bits L, and conceptually represented in a binary format with a sign bit s. This integer is represented using the number of leading zeros Z from a predetermined maximum bit number and the remaining bits r with a sign bit. This number of leading zeros Z is entropy coded based on the number of surrounding coefficients within a certain window that are not quantised to zero. For non zero integer coefficients the remaining bits r and sign bit s are coded as is.

As illustrated in Fig. 5, at the decoder the operation of the encoder is reversed (in as much as this is possible with quantisation). The quantised coefficients are inverse quantised 31. Finally an inverse Discrete Wavelet Transform is performed 32 on the resulting subbands to give the output image.

As noted previously, in the encoding process the discrete wavelet transform coefficients are quantised 22 to integer values. Let c represent a coefficient value and d its quantised values. Then the quantisation is performed as,

$$d = fix\left(\frac{c}{q}\right)$$

where q is a predetermined quantisation factor and fix is defined by,

$$fix(x) = \begin{cases} \lfloor x \rfloor & x \geq 0 \\ \lceil x \rceil & x \leq 0 \end{cases}$$

were $\lfloor \ \rfloor$ is the round down to nearest integer operator and $\lceil \ \rceil$ is the round up to nearest integer operator. At the encoder each coefficient in a subband is quantised to an integer value using this equation.

The inverse quantisation is given by,

$$c = q \times d + sign(d) \times \frac{q}{2}$$

where,

$$sign(d) = \begin{cases} -1 & d \le 0 \\ 0 & d = 0 \\ 1 & d \ge 0 \end{cases}$$

At the decoder each coefficient is inverse quantised using this inverse quantisation equation. The quantisation factor q can vary from subband to subband, or it can be fixed for the whole image. It can be coded in the header of the compressed image.

Coefficient Coding and Decoding

As shown in Fig. 6, each quantised coefficient is an integer value represented in a binary format with a sign bit. For the purpose of the description of the preferred embodiment, it is assumed with 15 bits and an extra sign bit (ie. L=16)). Thus

$$d = sign(d) \times b_{14} b_{13} \cdots b_0$$

where b_n is binary bit n. If the coefficient is non zero and the most significant bit number 26 is m, then:

$$b_{14} = b_{13} = \ldots = b_{m+1} = 0, b_m = 1,$$

and the number of leading zeros Z is,

$$Z = 14 - m,$$

If the coefficient is zero we set Z=15. The coefficient d is coded in two parts. First Z is entropy coded based on the context of the number of surrounding coefficients that are non-zero. A concise definition of surrounding coefficients is given below. Then for non-zero coefficients bits $b_{m-1}, \ldots, b_o$ and the sign bit are coded into the bit stream.

Surrounding Coefficient Context

Turning to Fig. 7, a subband eg. 35 is coded in raster scan order from top to bottom and left to right. If the current coefficient to be coded is marked with a cross 36,

the surrounding coefficients are considered to be the four surrounding coefficients indicated by the four empty squares 37-40. The surrounding coefficients are selected by a window with a shape as indicated in Fig. 7. If the cross in the window is aligned with the current coefficient the surrounding coefficients 37-30 are defined to be the coefficients that fall within the window.

The window illustrated follows a raster scan order. Hence, when the current coefficient is being decoded the surrounding coefficients have already been decoded, and thus the decoder knows whether or not they are non zero. The context for the current pixel is determined by the number of surrounding coefficients 37-70 that are non-zero. In this case there are 5 contexts corresponding to 0, 1, 2, 3 or 4 surrounding pixels that are non-zero. For the coefficients in the first row or column a modified window is used that includes only coefficients in the current subband.

Obviously different windows can be used. Ideally, the surrounding coefficients must come before the current coefficient in the raster scan order. This is so the decoder knows their value before decoding the current pixel. The four coefficient window of Fig. 7 has been selected for a good compromise between complexity, which grows with the number of contexts, and compression efficiency, which increases with the number of contexts, at least up to a certain point.

Context Entropy Coding

As noted previously, the number of leading zeros Z of each coefficient is coded with a context based entropy coder. Preferably, this is a standard arithmetic coder. Arithmetic coding is described in Witten et. al., "Arithmetic coding for data compression", Communications of the ACM, Volume 30, No. 6, June 1987. Preferably fixed histograms are utilised. The histograms are then coded into the compressed image header and are used by the decoder.

In the preferable implementation the arithmetic coder

is initialised with different histograms for each subband. For a given subband a histogram with 16 bins is generated for each context as illustrated in Fig. 8. For a given context the i^{th} bin of the histogram is the count of the number of coefficients with i leading zeros and whose context of surrounding coefficients is the given context. For example, for a context of 2 surrounding non-zero coefficients, bin 7 of the histogram is the count of the number of coefficients with 7 leading zeros and who have 2 surrounding non-zero pixels. These histograms are made prior to the coding process. As an alternative an adaptive arithmetic coder could be utilised.

At the decoder these histograms are used to decode the leading zero information for the subband coefficients, which is contained in the coded bit stream. For the non-zero coefficients, the remaining bits and sign bit are read from the coded bit stream and hence the quantised coefficients can be constructed.

The preferred embodiment also has application to other image formats. For example, full colour images can be encoded via separate colour channels or the usual chrominance compression techniques as utilised in the JPEG standard can be applied so as to produce reduced chrominance data.

Additionally, the principles of the preferred embodiment can be equally extended to other forms of data such as sound data etc. and the preferred embodiment has application wherever wavelet transforms are suitable. Additionally, the preferred embodiment can be applied to other forms of transformed data for example, the discrete cosine transform process in addition to the wavelet packet and cosine packet transform techniques as described in the aforementioned survey article.

The preferred embodiment is ideally implemented on a suitably programmed computer device.

It would be appreciated by a person skilled in the

art that numerous variations and/or modifications may be made to the present invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to be considered in all respects to be illustrative and not restrictive.

We Claim:

1. A method of compressing data comprising:

applying a transform to the data to produce transformed data having a series of parts;

entropy encoding the magnitude of the transformed data of at least one of said parts based on the magnitude of the surrounding transformed data; and

separately encoding the value of said transformed data.

2. A method as set out in paragraph 1 wherein said entropy encoding utilizes the number of non-zero coefficients surrounding a spatial location of a corresponding transformed data value.

3. A method as set out in paragraph 2 wherein the surrounding values utilized are previously entropy encoded values adjacent to a current spatial location of the corresponding transformed data value.

4. A method as set out in paragraph 1 wherein said transforming step further comprises quantizing transformed portions of said data to integer values.

5. A method as set out in paragraph 4 wherein said integer values include a sign bit and a predetermined number of coefficient bits.

6. A method as set out in any previous paragraph wherein said transform comprises wavelet transforming the data.

7. A method as set out in paragraph 6 wherein said parts comprise each of the sub-band components of the wavelet transform which are separately entropy encoded.

8. A method as set out in paragraphs 6 or 7 wherein the lowest frequency sub-band component is separately encoded.

9. A method as set out in any previous paragraph wherein said data comprises image data describing an image.

10. A method as set out in any previous paragraph wherein said magnitude encoding comprises encoding the

number of leading zeros in transformed data values.

11. An apparatus when implementing the method as set out in any of paragraphs 1 to 10.

Abstract

A method of compressing data is disclosed comprising applying a transform to the data to produce transformed data having a series of parts; entropy encoding the magnitude of the transformed data of at least one of said parts; and separately encoding the value of said transformed data. Preferably said entropy encoding utilizes the number of non-zero coefficients surrounding a spatial location of a corresponding transformed data value and the entropy encoding can comprise encoding the number of leading zeros in transformed data values. The method further comprises quantizing transformed portions of said data to integer values including a sign bit and a predetermined number of coefficient bits. Ideally, the preferred embodiment includes wavelet transforming the data with each of the sub-band components of the wavelet transform being separately entropy encoded. The present invention is ideally suited to the compression of image data.

DATED this TWENTY-NINTH day of SEPTEMBER 1997

Canon Kabushiki Kaisha

Canon Information Systems Research Australia Pty Ltd

Patent Attorneys for the Applicants
SPRUSON & FERGUSON

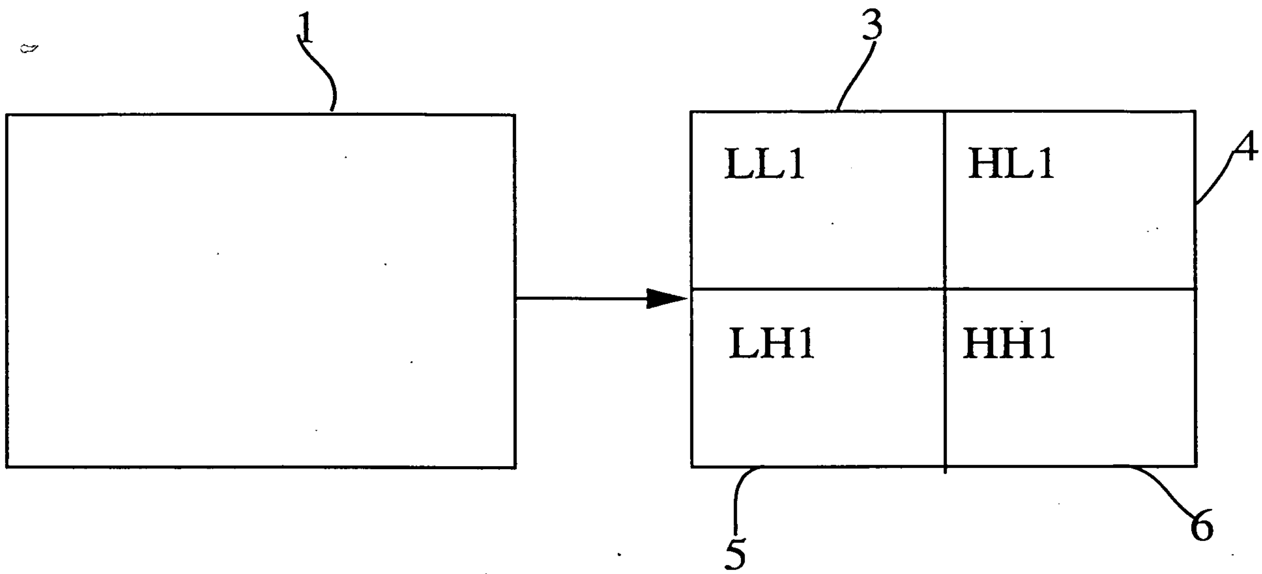


Fig 1

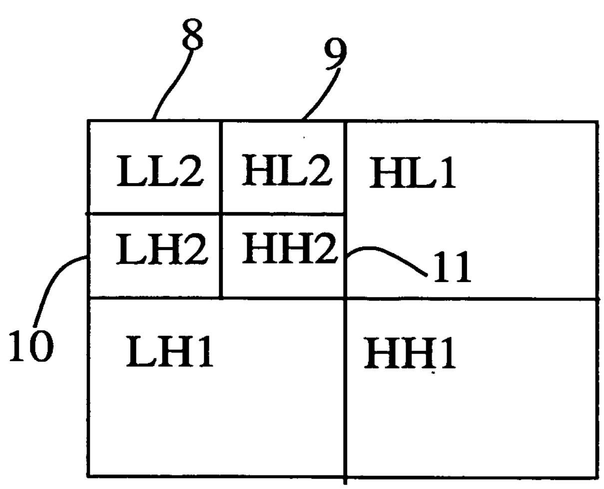


Fig 2

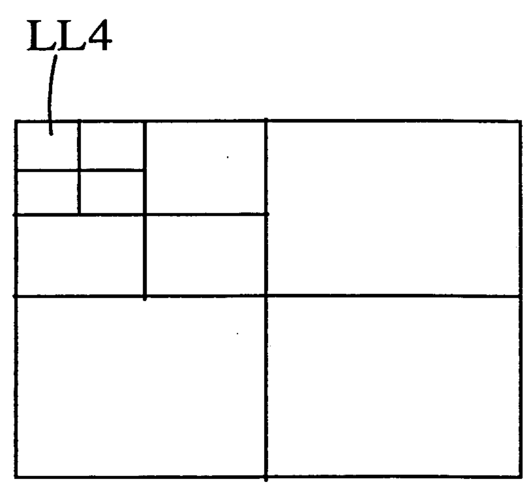
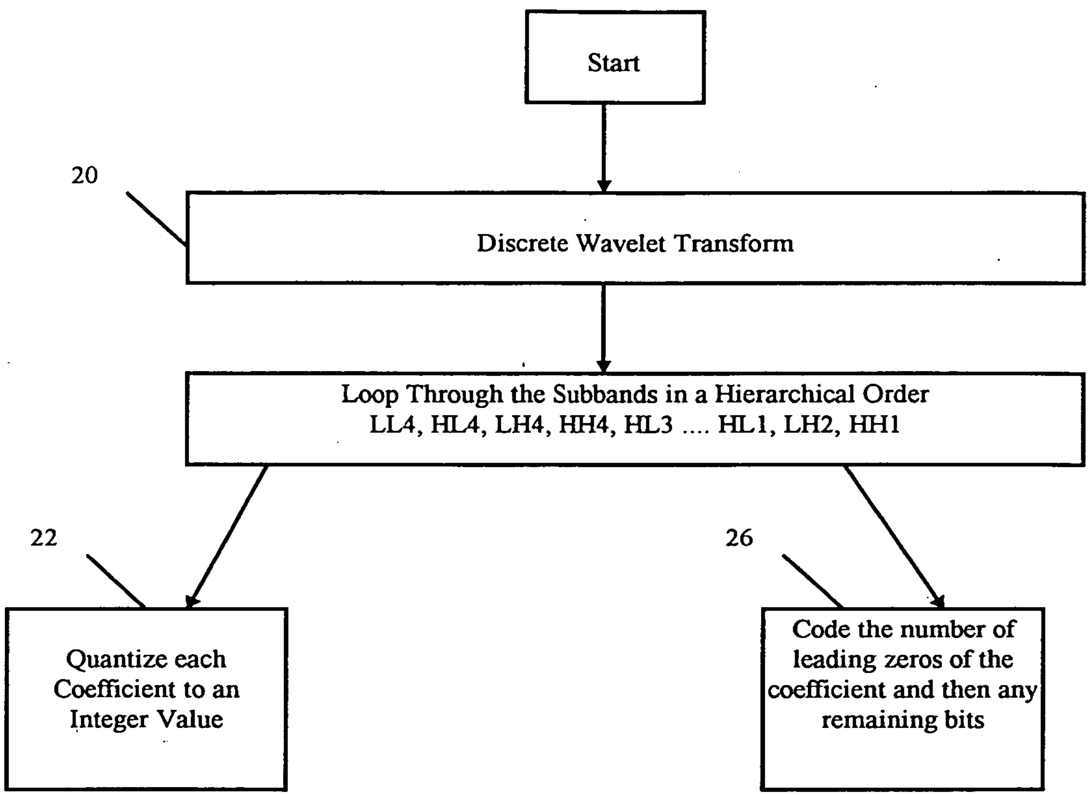
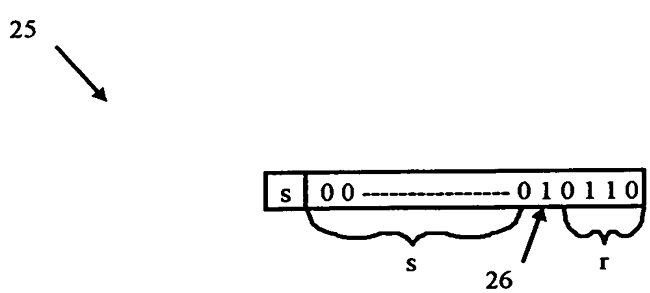


Fig 3



Start
20
Discrete Wavelet Transform
Loop Through the Subbands in a Hierarchical Order
LL4, HL4, LH4, HH4, HL3 .... HL1, LH2, HH1
22
26
Quantize each Coefficient to an Integer Value
Code the number of leading zeros of the coefficient and then any remaining bits

Fig. 4



25
s
00 ---------------- 010110
s
26
r

Fig. 6

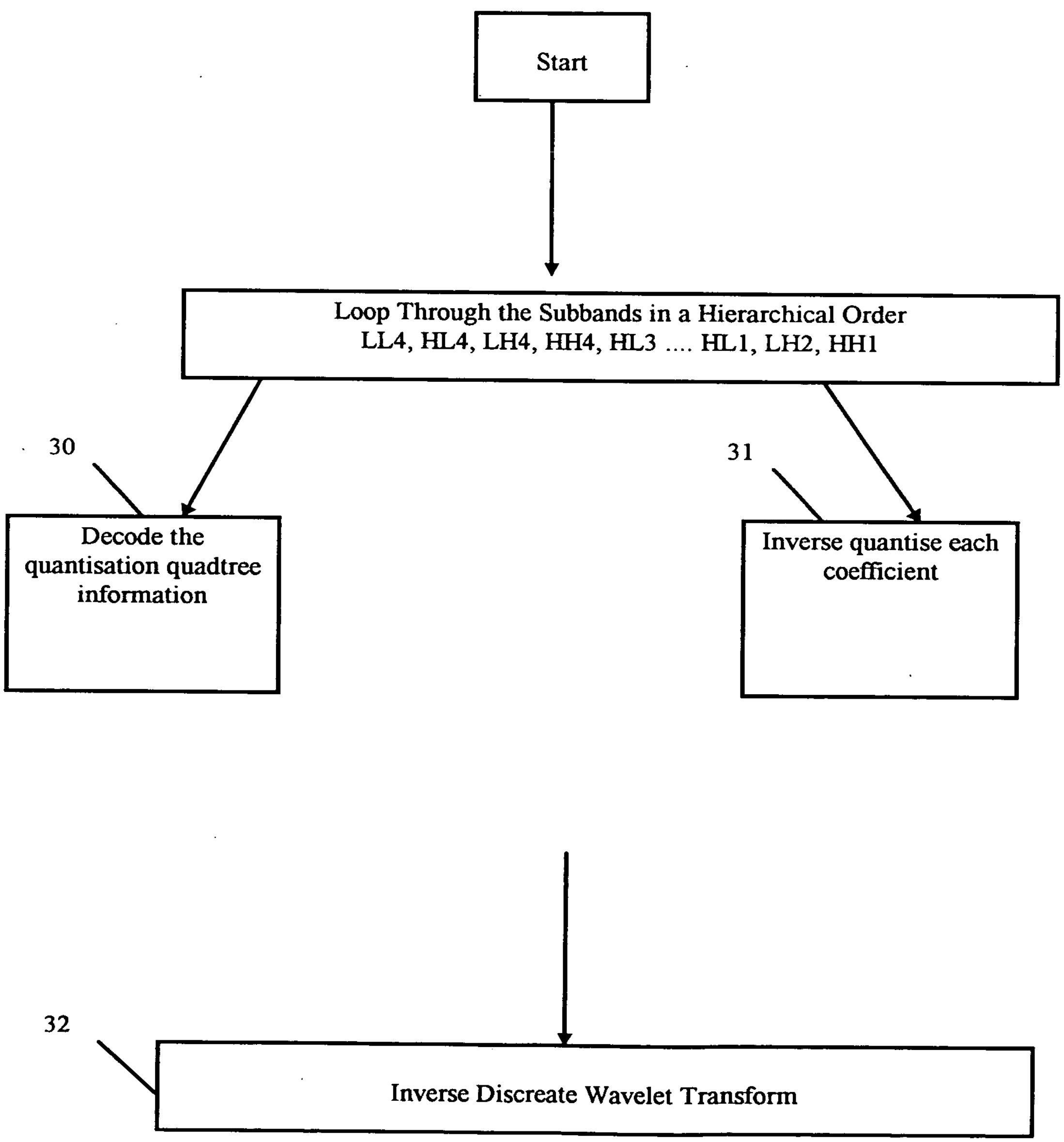

Start
Loop Through the Subbands in a Hierarchical Order
LL4, HL4, LH4, HH4, HL3 .... HL1, LH2, HH1
30
Decode the quantisation quadtree information
31
Inverse quantise each coefficient
32
Inverse Discreate Wavelet Transform

Fig. 5

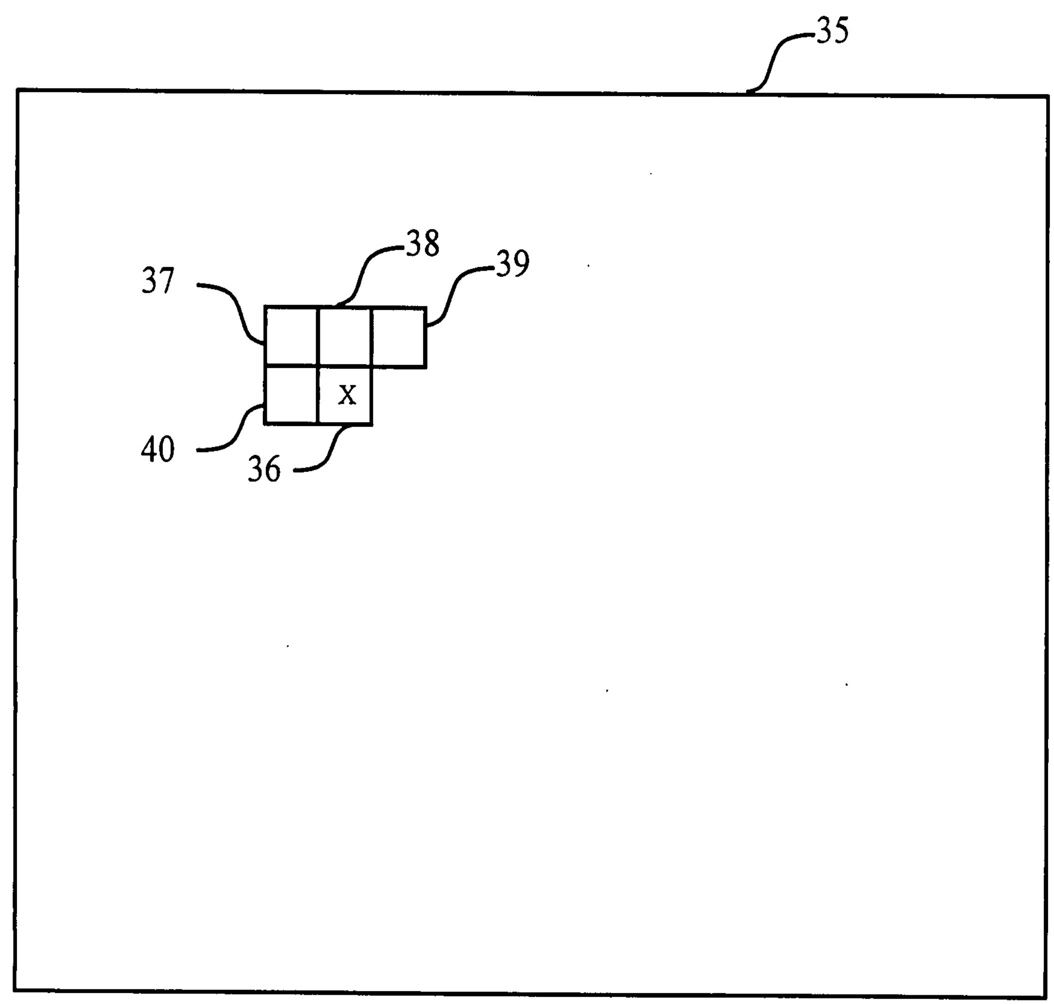


Fig. 7

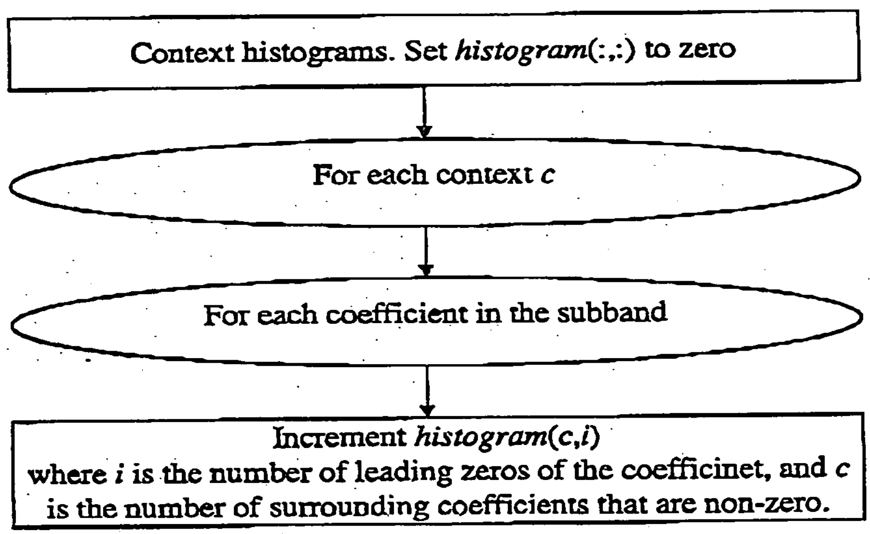

Context histograms. Set *histogram*(:,:) to zero
For each context *c*
For each coefficient in the subband
Increment *histogram*(*c*,*i*)
where *i* is the number of leading zeros of the coefficinet, and *c* is the number of surrounding coefficients that are non-zero.

Fig. 8.